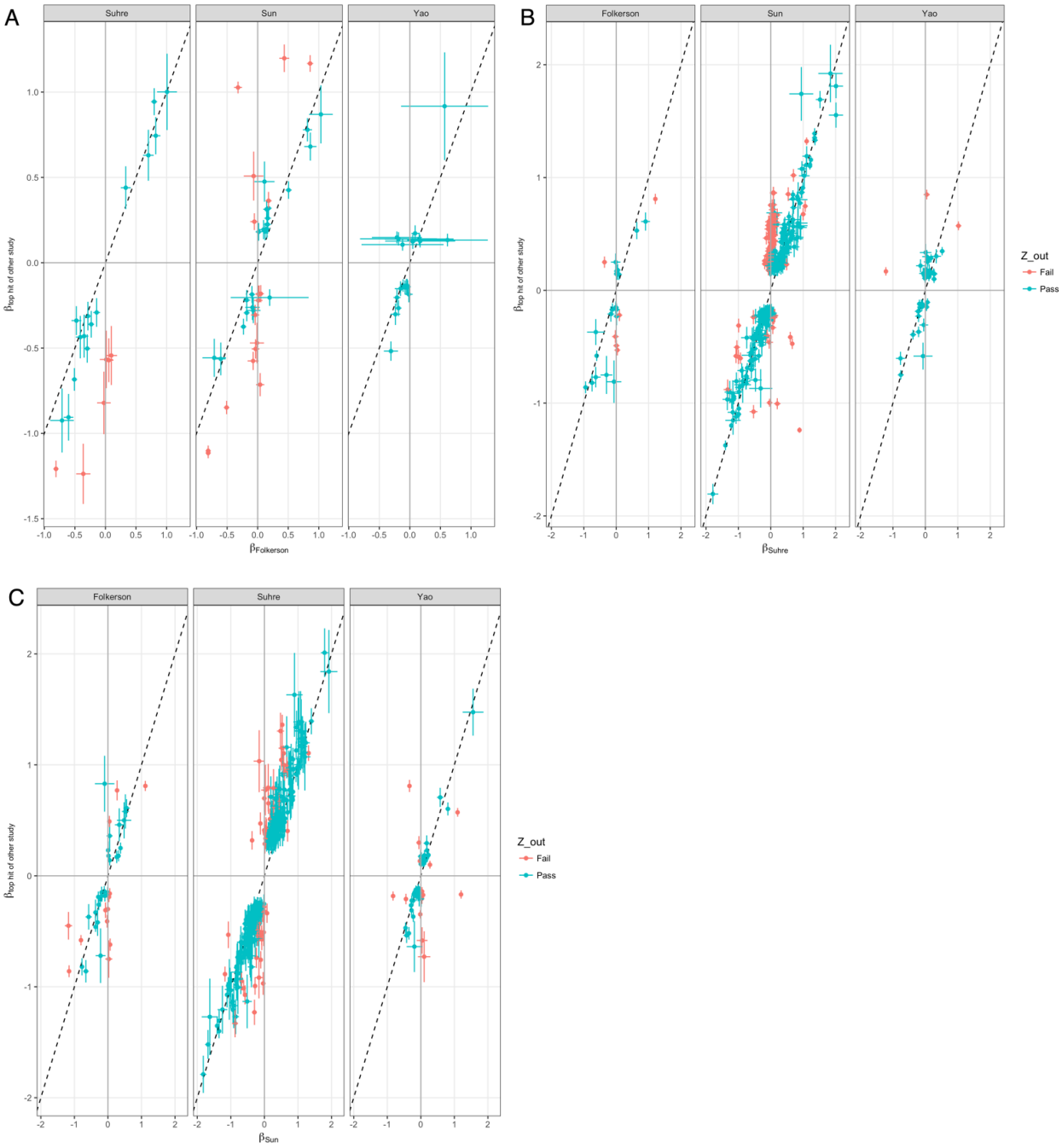
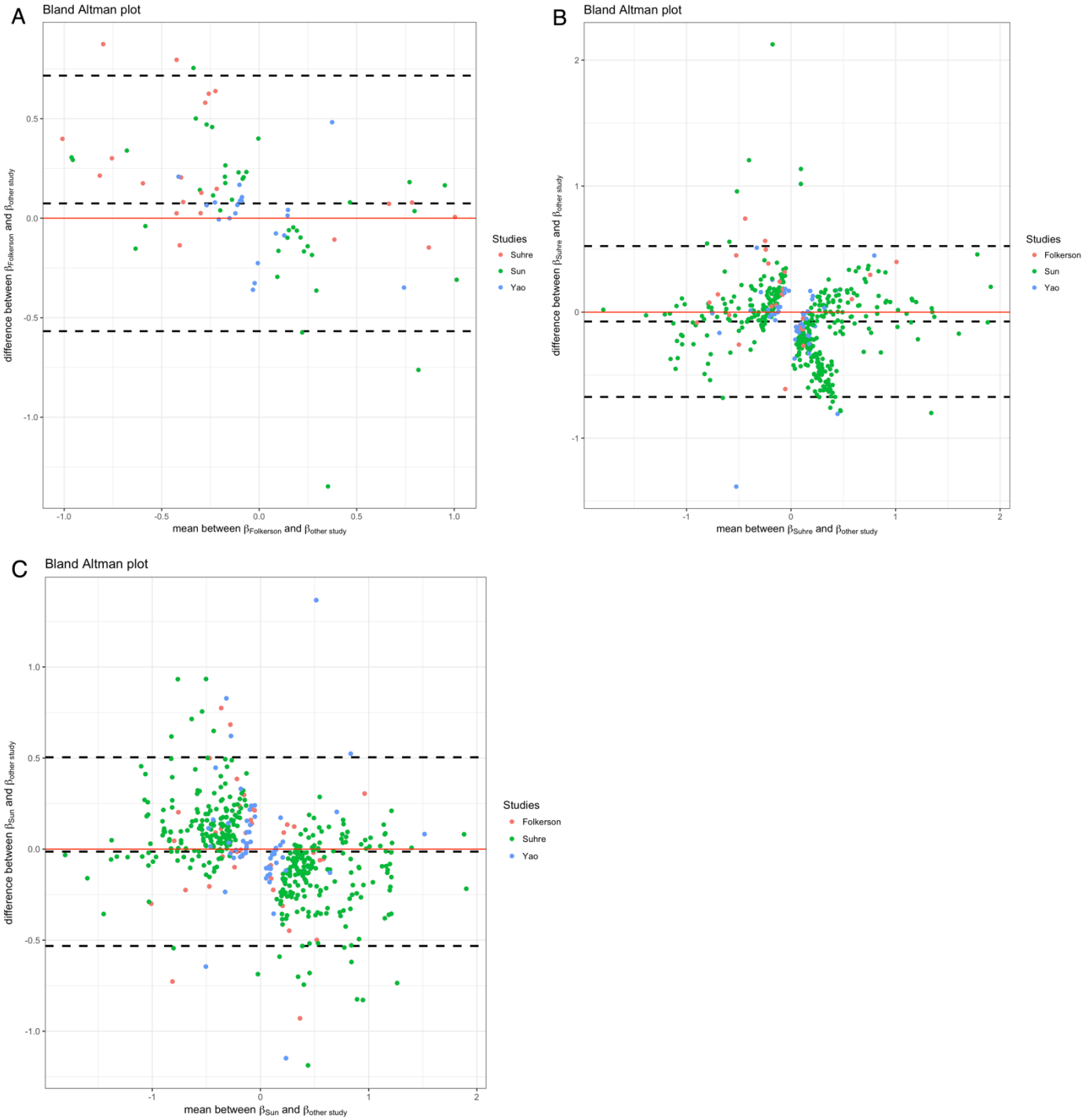
Supplementary documents

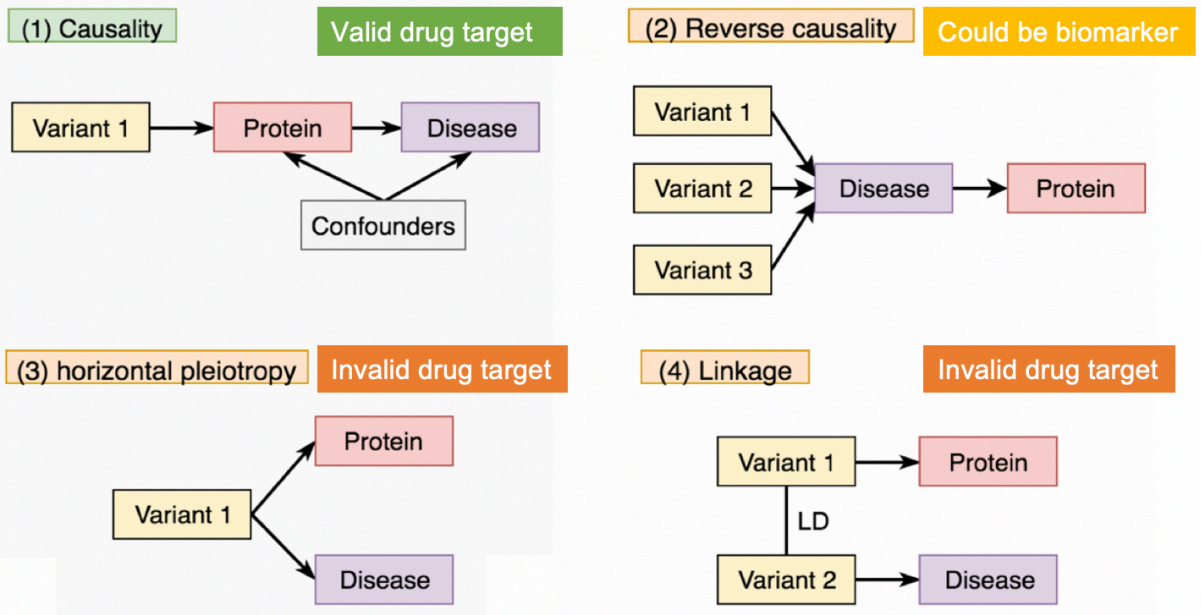
Supplementary Figures



Supplementary Figure 1. Scatter plot showing the pair-wise instrument validation across the pQTL GWAS used in the MR study. A) SNP effects from Folkersen vs SNP effects from Sun et al, Suhre et al and Yao et al; B) SNP effects from Suhre vs SNP effects from Sun et al, Folkersen et al and Yao et al; C) SNP effects from Sun vs SNP effects from Folkersen et al, Suhre et al and Yao et al. Notation: X-axis refers to the SNP effects from the lookup study; Y-axis refers to the SNP effects from the other studies; dotted line is the identity line. Each dot refers to one SNP in the comparison, dots in green refer to the SNP effects which have a pair-wise Z score > 5, whereas the dots in red refer to SNP effects which have a pair-wise Z score < 5.



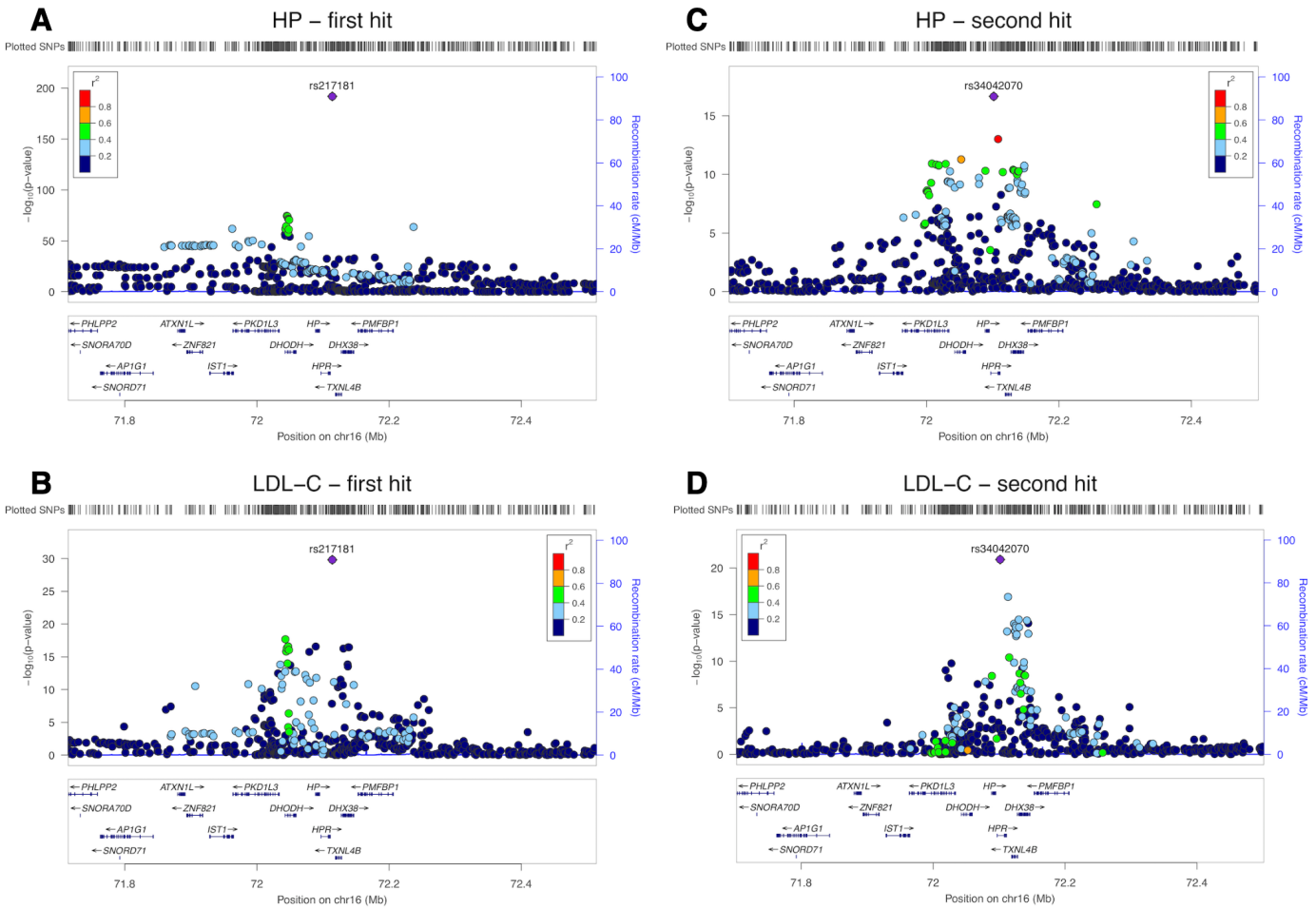
Supplementary Figure 2. Bland-Altman plots showing pair-wise instrument validation across the pQTL GWASs used in the MR study. A) SNP effects from Folkersen vs SNP effects from the other three studies; B) SNP effects from Suhre vs SNP effects from the other three studies; C) SNP effects from Sun vs SNP effects from the other three studies. Notation: X-axis refers to the mean across effects; Y-axis refers to the difference between effects. The three dotted lines refer to the central estimate and the 95% confidence interval lines of the Bland-Altman test. Each dot refers to one SNP in the comparison, dots in red, green and blue refer to the SNP effects from each of the different studies.



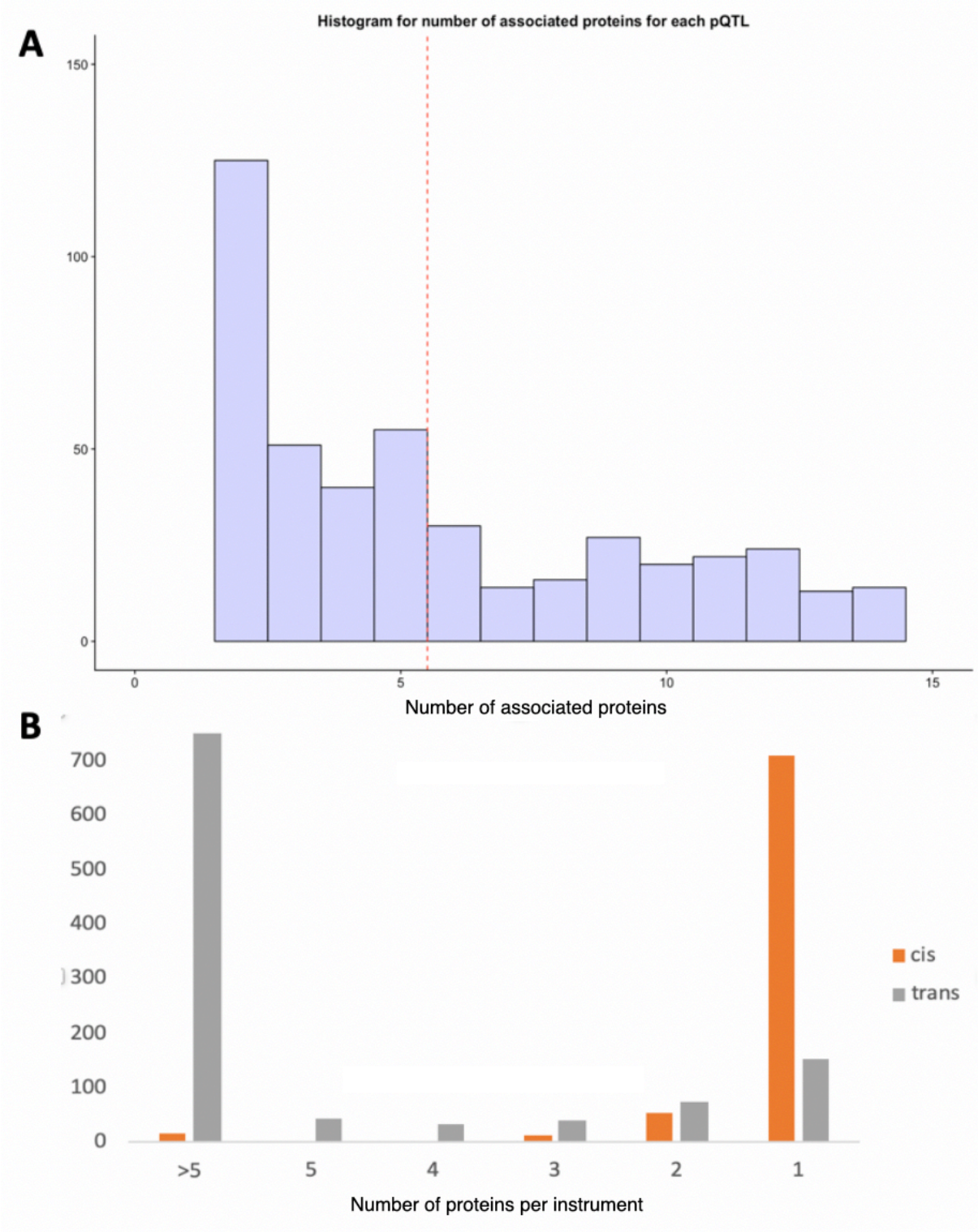
Supplementary Figure 3. MR and colocalization models. Model 1 – Causality: A genetic variant affects disease risk by changing protein levels; Model 2: Reverse causality Genetic variants affect disease risk through pathways other than via the protein of interest. The disease has a downstream effect on protein levels; Model 3 – Horizontal pleiotropy: a genetic variant influences both protein levels and disease risk by two independent biological pathways; Model 4 – confounding by LD: a genetic variant (variant 1) that influences protein levels is correlated with a second variant (variant 2) that influences disease risk. Colocalization analysis can distinguish Model 4 from Model 1 or Model 3.

		IL23R		
		Marginal SNP effects	Condition on the 2nd pQTL	Condition on the 1st pQTL
Crohn's disease	Marginal SNP effects	0%	0%	0%
	Condition on the 1st CD signal	0%	0%	62.90%
	Condition on the 2nd CD signal	0%	99.30%	0%
	Condition on the 3rd CD signal	0%	0%	0%
	Condition on the 4th CD signal	0%	0%	0%

Supplementary Figure 4. Heatmap of the colocalization evidence for IL23R association on Crohn's disease (CD) in the *IL23R* region. The 15 cells refer to the 15 pair-wise combinations of pair-wise conditional and colocalization analysis. The three columns refer to the SNP effects of IL23R protein level used in the colocalization analysis (marginal SNP effect, joint SNP effect after conditioning on the 2nd IL23R signal (rs3762318) and the joint SNP effect after conditioning on the 1st IL23R signal (rs11581607)). The five rows refer to the SNP effects of Crohn's disease used in the colocalization analysis (marginal SNP effect, joint SNP effects after conditioning on the 1st, 2nd, 3rd and 4th CD signals). The darker red colour refers to stronger colocalization evidence.

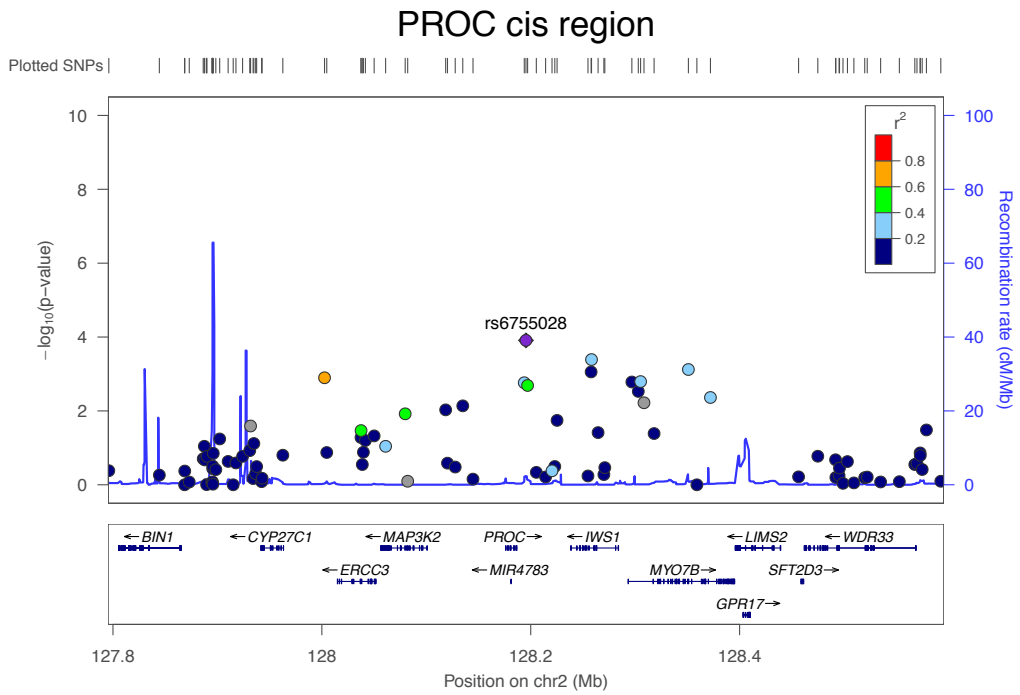


Supplementary Figure 5. Regional association plot showing multiple association peaks for Haptoglobin (HP) and LDL cholesterol in the cis region. The second independent hit for LDL cholesterol colocalised with the top hit for HP (rs217181) after conditioning on the top hit for LDL cholesterol (rs2000999) (colocalization probability = 99.9%). (A) HP after conditioning on the second pQTL HP rs34042070; (B) LDL-C after conditioning on the second pQTL HP rs34042070; (C) HP after conditioning on the top pQTL HP rs217181; (D) LDL-C after conditioning on the top pQTL HP rs217181.The HP data are from Sun et al. and the LDL cholesterol data are from GLGC consortium.

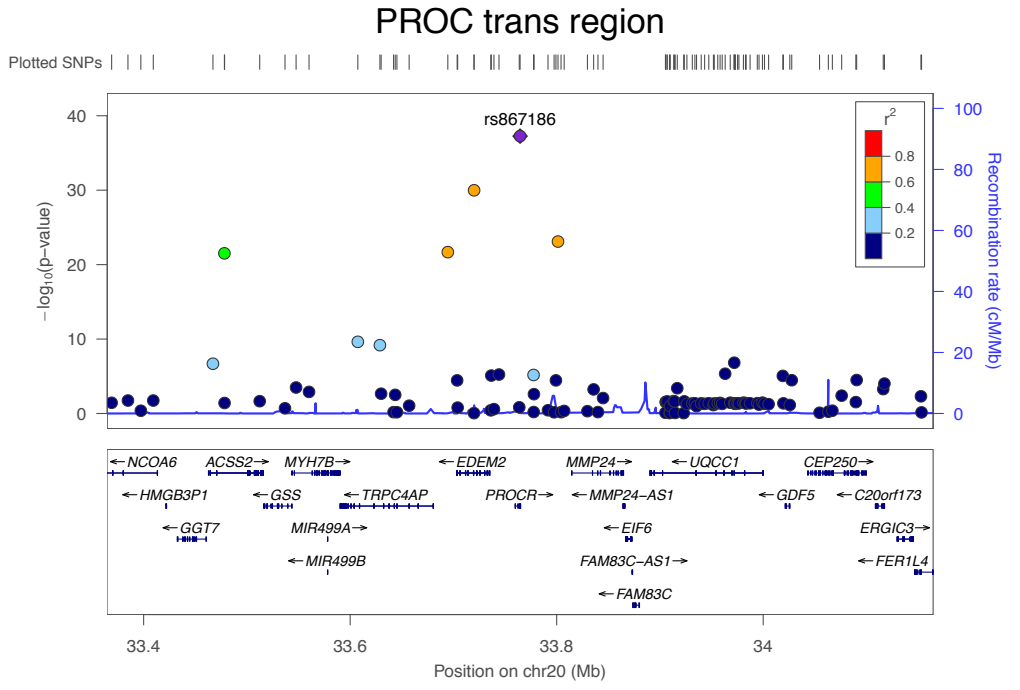


Supplementary Figure 6. Histogram representing the distribution of the number of proteins associated with each pQTL. (A) a zoomed in histogram which only showed pQTLs associated with fewer than 15 proteins. (B) a histogram of number of proteins each pQTL associated with, the grey bar refers to trans pQTL; the orange bar refers to cis pQTL. There is a clear trend that the trans pQTLs were associated with more proteins (potentially pleiotropic) compared to the cis pQTLs.

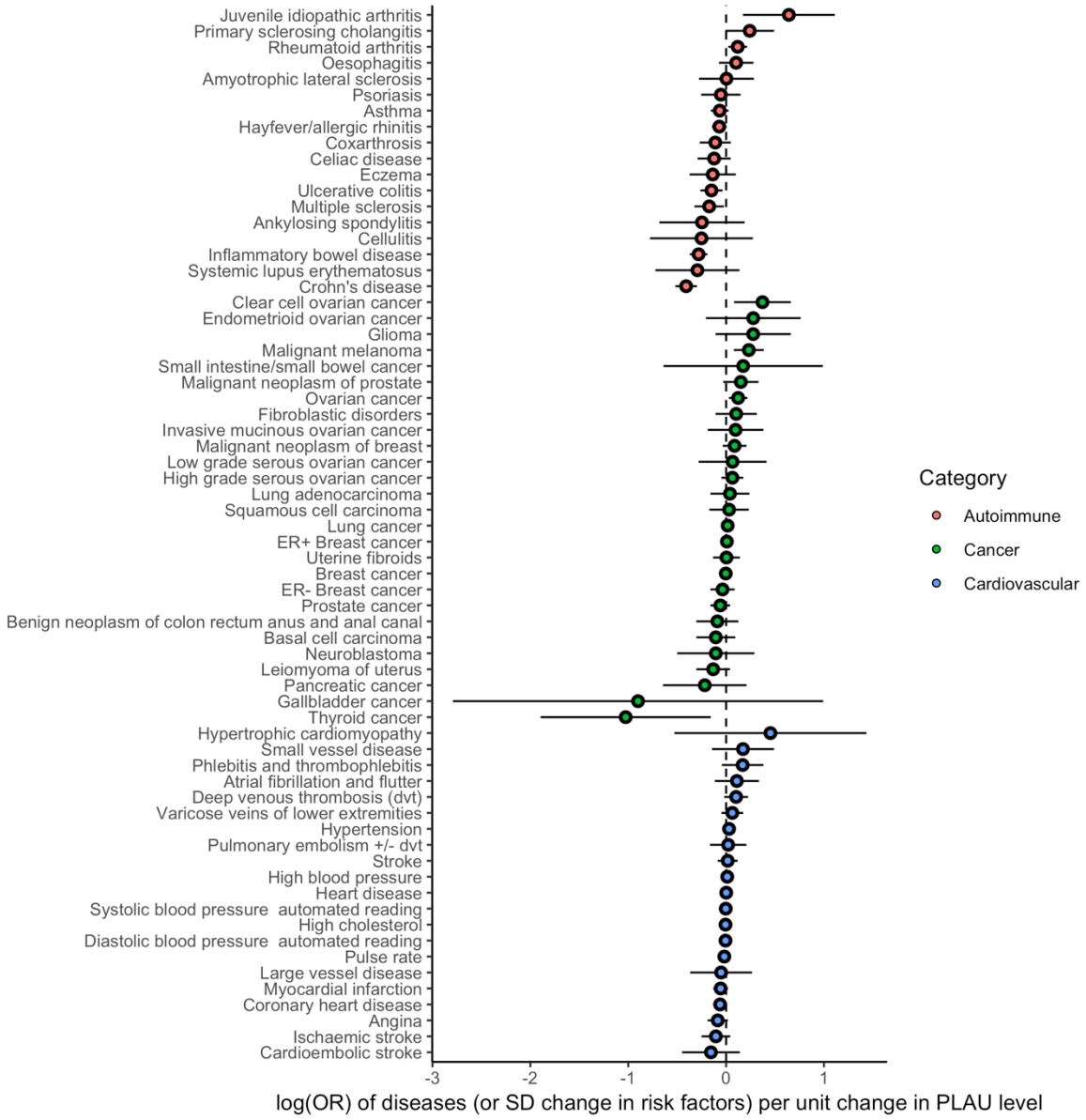
A



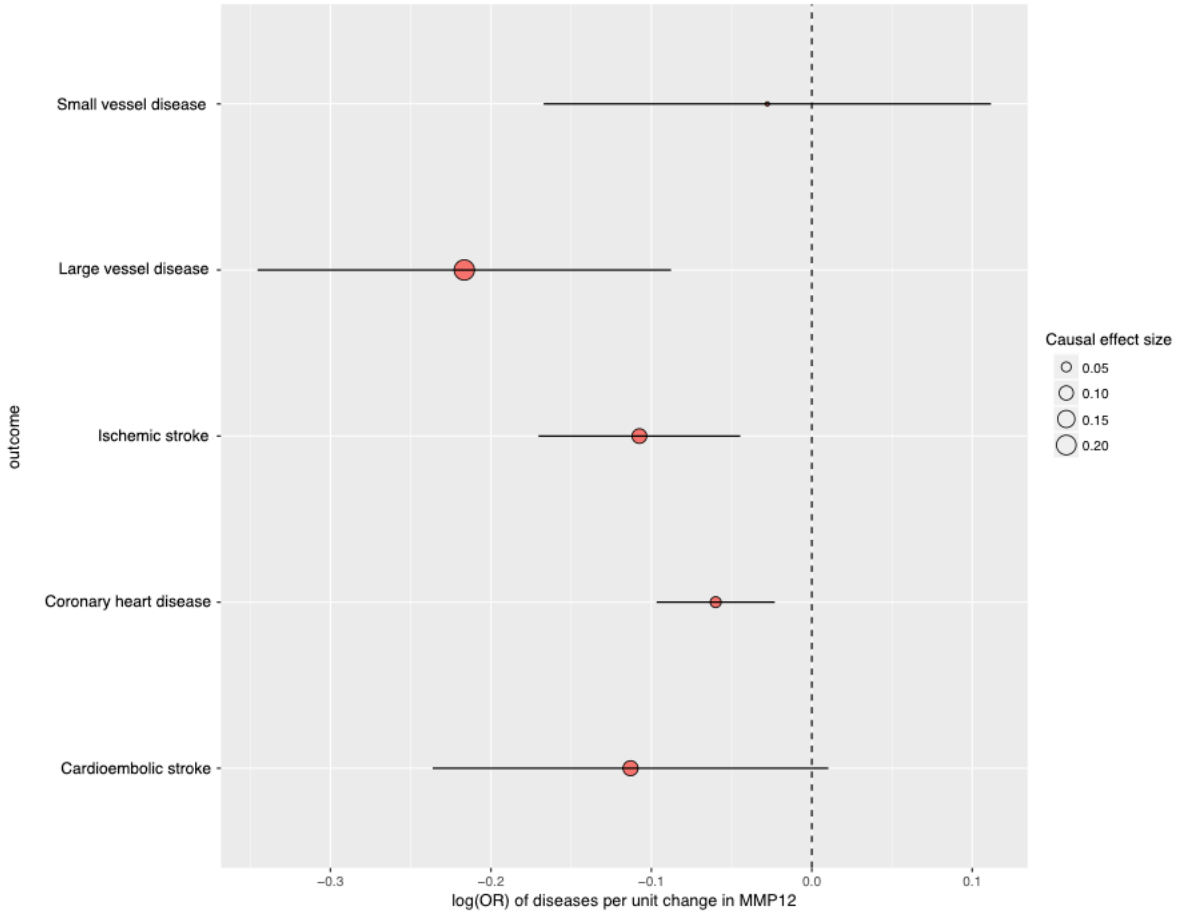
B



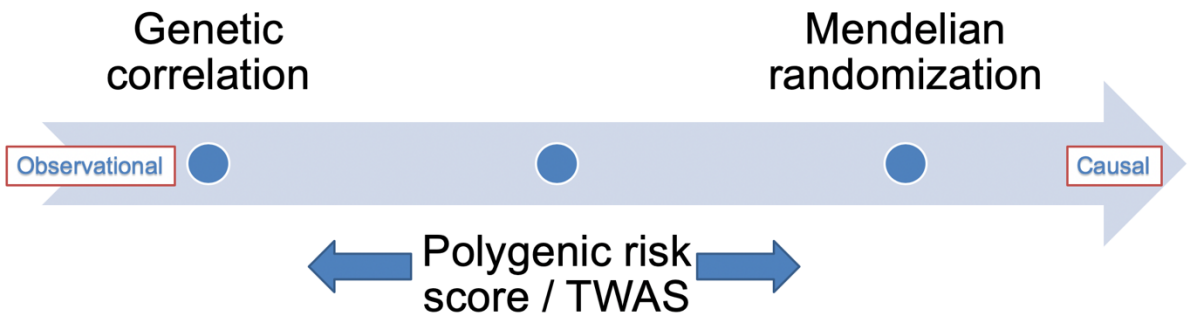
Supplementary Figure 7. Regional association plot of Protein C (PROC) protein levels in two regions. (A) there is little evidence that SNPs are associated with PROC protein levels in the cis region. (B) there is strong evidence that SNPs are associated with PROC in the trans *PROCR* region. The PROC data is from Suhre et al.



Supplementary Figure 8. Forest plot of MR estimates for plasma PLAU levels on immune-mediated phenotypes, cardiovascular phenotypes and cancers. The X-axis refers the log odds ratio of disease (or SD change in risk factor) per unit change in plasma PLAU level. The Y-axis refers to phenotypes tested in this Phenome-wide MR. The error bar refers to the 95% confidence interval of the causal estimates. Colours refer to different categories of phenotypes.



Supplementary Figure 9. Forest plot showing MMP12 MR associations on coronary heart disease and stroke. The X-axis refers the log odds ratio of disease per unit higher MMP12 level. The Y-axis refers to different subtypes of cardiovascular diseases. The error bar refers to the 95% confidence interval of the causal estimates.



Supplementary Figure 10. Difference between Mendelian randomization, polygenic risk score / TWAS and genetic correlation. Note: Genetic correlation estimates the overall genetic overlap of SNPs across the whole genome, which is closer to observational correlation between two phenotypes with no direction. Mendelian randomization estimates the causal relationship between two phenotypes with inference of direction, typically using SNPs robustly associated with the exposure (e.g. protein level) as instruments, where the instruments can be in either cis and/or trans regions. The SNP effect on human phenotype will be estimated first and then meta-analysed using various models such as inverse variance weighted (IVW) meta-analysis. The polygenic risk score association and TWAS are similar, and both use individual level genotype data. The method contains two steps, 1) using SNPs within the cis region to predict the expression of a gene; 2) correlate the predicted expression of the gene with the human phenotype.

Supplementary Notes

Protein-trait associations in three disease areas

We found that our MR findings were clustered in three areas that have not been described well in the MR literature: blood pressure (AGT, ADM, ERAP2, FN1, SWAP70, CXCL16 and IGFBP3), lung function (ADAM19, APOF, GPC5, SERPINF1, MFAP2) and immune mediated disease (IL23R, IL6R, IL18R1, FCRL3, ICAM5 and PLAU).

For blood pressure, Adrenomedullin (ADM) is a hormone involved in vascular tone, and is a well-studied target for blood pressure [1 2] *Endoplasmic Reticulum Aminopeptidase 2* (ERAP2) has been linked to pre-eclampsia [3], for which there are currently no effective drugs. Rare mutations in *Fibronectin 1* (FN1) have been associated with glomerulopathy with fibronectin deposits [4], which leads to hypertension [5]. *C-X-C Motif Chemokine Ligand 16* (CXCL16) is an interferon-γ-regulated chemokine and scavenger receptor for oxidized low-density lipoprotein that is expressed in atherosclerotic lesions [6] and has been linked to Chronic Kidney Disease [7]. These connections may suggest the association of CXCL16 on DBP is at least in part, through the atherosclerosis or renal function. Insulin-like growth factor binding protein 3 (IGFBP3) levels have been previously reported to associate with hypertension [8], and a SNP in IGFBP3 has been found to associate with increased long-term average pulse blood pressures [9].

For lung function, *A Disintegrin And Metalloproteinase Domain 19* (ADAM19) is a metalloproteinase, which may have a role in tissue remodelling [10]. *Apolipoprotein F* (APOF) is a lipoprotein, which may associate with chronic obstructive pulmonary disease (COPD) [11]. *Glypican Proteoglycan 5* (GPC5) is a cell surface heparin sulfate proteoglycan) that has been linked to lung cancer [12]. *Serpin Family F Member 1* (SERPINF1) inhibits angiogenesis [13] and rare mutations are associated with osteogenesis imperfecta [14], but a link to lung function is unclear. *Microfibril Associated Protein 2* (MFAP2) is an antigen of elastin associated fibrils so may be important in tissue remodelling in the lung. However, the MFAP2 locus was associated with both height and lung function [15]. It is possible that the association we identified could be mediated by height.

For the immune mediated traits, an existing IL23R antagonist, ustekinumab, demonstrated efficacy in reducing Crohn's diseases in a recent phase III clinical trial (www.clinicaltrials.gov) [16]. Our MR analysis further linked IL23R inhibition with psoriasis. IL18R1 is a known eczema locus which replicated in both the Japanese and European population [17 18]. Our MR study suggested that IL18R1 could be an effector / causal gene for eczema. Polymorphisms within *Fc Receptor Like 3* (FCRL3) have been found to associate with rheumatoid arthritis in the Chinese population [19]. A link between *Intercellular Adhesion Molecule 5* (ICAM5) and Crohn's disease is not yet clear in the literature.

MR results replicated previous findings

Some of our MR results replicated those previously described by others. For example, Sun *et al* suggested that the TNFRSF11A associated variant rs884205 was also associated with Paget's disease [20]. Our Wald ratio analysis confirmed the positive association between this pQTL and Paget's disease (OR=8.56, 95%CI=4.78 to 15.31, P= 4.72×10^{-13}). The colocalization analysis further confirmed that the two traits share the same casual variant within the TNFRSF11A region (PP=99%).

In addition, we replicated the apparent effect of MMP12 pQTL on CHD (OR=0.94, 95%CI=0.91 to 0.98, P= 0.0014) and stroke reported in Sun *et al* [20]. We extended this analysis to stroke subtypes and found that MMP12 pQTL were associated with ischemic stroke (OR=0.90, 95%CI=0.84 to 0.96, P=0.0008) and large vessel disease (OR=0.81, 95%CI=0.71 to 0.92, P=0.00098) but not strongly with cardioembolic stroke (OR=0.89, 95%CI=0.79 to 1.01, P=0.07) and not with small vessel disease (OR=0.97, 95%CI=0.85 to 1.12, P=0.70) (**Supplementary Table 25**, **Supplementary Figure 9**).

Bi-directional MR and Steiger filtering results

For the bi-directional MR, we modelled complex traits as our exposure (data from MR-Base) and plasma protein level as our outcome (full summary statistics of proteins were available for Sun *et al* and Folkersen *et al* [20,21]). For clarity – this analysis does not necessarily implicate the disease status as being causal for the protein levels, but it indicates that genetic liability to the disease may influence protein levels. In total, the relationship between 104 diseases and 206 proteins were tested (841 tests, Bonferroni-adjusted P = 5.9×10^{-5}). We found no strong evidence of reverse causality between protein level and disease for the majority of protein-trait associations (**Supplementary Data 1**). However, there were exceptions for traits with a strong genetic signal at the *APOE* locus.

Due to a lack of full summary statistics for some pQTL studies, we were not able to conduct bi-directional MR for all MR findings. Instead, we applied Steiger filtering as an alternative method to infer the causal direction for our MR. For MR findings using multiple instruments, the Steiger filtering tested the directionality of each instrument on exposure and outcome separately (rather than testing the overall directionality between exposure and outcome). Among all 397 protein-trait associations we tested, 360 had enough statistical power for the Steiger filtering analysis (to detect a nominally significant P value <0.05). For these 360 cases, 303 (84.2%) showed evidence that the direction of the association is from protein to human traits (**Supplementary Table 7, 8, 11, 12 and 13**). We found a substantially more reverse causal instances for trans (31.2%) than there are for cis (0.6%).

	Trans	Cis
SF-FALSE	53	1
SF-TRUE	117	178
All	170	179
Percentage	31.2%	0.6%

Note: SF-FALSE refers to protein-phenotype association with evidence of reverse causality using Steiger filtering; SF-TRUE means the association with no evidence of reverse causality.

Case study for drug repurposing

As an example of drug repositioning, our phenome-wide MR analysis suggested that lifelong higher urokinase-type plasminogen activator (PLAU) levels are associated with lower inflammatory bowel disease (IBD) risk (OR=0.75, 95%CI= 0.69 to 0.83, P= 1.28×10^{-9}; **Supplementary Figure 8**), potentially identifying a repositioning opportunity for IBD. However, we note this opportunity with caution given the multitude of considerations in such a strategy. For example, the drug Kinlytic (urokinase) was initially developed for use as a thrombolytic in the treatment of acute myocardial infarction and ischaemic stroke, and thus a target-mediated adverse effect is an increase in bleeding and potential haemorrhage. In addition, the current agent is administered intravenously which precludes it as an option as a long-term preventative treatment. While our data suggest that Kinlytic might be protective in the aetiology of IBD, a careful risk benefit assessment would be required as part of an investigation into whether drugs targeting urokinase might be repurposed for the treatment of IBD.

Description of ALSPAC study

Pregnant women resident in Avon, UK with expected dates of delivery 1st April 1991 to 31st December 1992 were invited to take part in the study. The initial number of pregnancies enrolled is 14,541 (for these at least one questionnaire has been returned or a "Children in Focus" clinic had been attended by 19/07/99). Of these initial pregnancies, there was a total of 14,676 foetuses, resulting in 14,062 live births and 13,988 children who were alive at 1 year of age.

When the oldest children were approximately 7 years of age, an attempt was made to bolster the initial sample with eligible cases who had failed to join the study originally. As a result, when considering variables collected from the age of seven onwards (and potentially abstracted from obstetric notes) there are data available for more than the 14,541 pregnancies mentioned above.

The number of **new pregnancies** not in the initial sample (known as Phase I enrolment) that are currently represented on the built files and reflecting enrolment status at the age of 24 is 904 (452, 254 and 198 recruited during Phases II, III and IV respectively), resulting in an additional 811 children being enrolled. The phases of enrolment are described in more detail in the cohort profile paper (see footnote 4 below). Please note that phase 4 enrolment (age 18-24) is not currently included in the cohort profile.

The total sample size for analyses using any data collected after the age of seven is therefore 15,247 pregnancies, resulting in 15,458 foetuses. Of this **total sample** of 15,656 foetuses, 14,973 were **live births** and 14,899 were **alive at 1 year of age**.

A 10% sample of the ALSPAC cohort, known as the **Children in Focus (CiF) group**, attended clinics at the University of Bristol at various time intervals between 4 to 61 months of age. The CiF group were chosen at random from the last 6 months of ALSPAC births (1432 families attended at least one clinic). Excluded were those mothers who had moved out of the area or were lost to follow-up, and those partaking in another study of infant development in Avon.

Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees.

Please note that the study website contains details of all the data that is available through a fully searchable data dictionary and variable search tool: http://www.bristol.ac.uk/alspac/researchers/our-data/

The protocol of the instrument validation

Because instruments were identified from five independent studies performed using different analytical platforms (Box 1), we developed a protocol for instrument validation. **Figure 1** summarises the 2 key analyses used for instrument validation.

Study	Platform	Sample size	Number of proteins (with pQTLs)	Number of instruments (pQTLs)
Sun *et al*	SOMAScan	3301	1478	1981
Emilsson *et al*	SOMAScan	3200	776	875
Suhre *et al*	SOMAScan	1000	284	539
Folkersen *et al*	Olink	3394	58	80
Yao *et al*	xMAP	6861	60	131

Box 1. The study level information of the 5 pQTLs studies.

1 Harmonisation of Protein IDs and instruments

Since the 5 previous studies measured plasma protein levels using different probes from three different platforms (SOMAScan, Olink and Luminex xMAP), we mapped the platform ID for each protein analyte from each study to Uniprot IDs (and associated annotations) based on annotations provided by the platform vendors and manual review. We then grouped the analytes based on their Uniprot IDs. The end product included columns corresponding to the 'platform ID' of each the analyte from each study, the full protein name, gene symbol, Uniprot ID, Ensembl gene id and GRch38 gene location (**Supplementary Table 26**). Since a single probe from any platform can map to multiple proteins, each of the protein names, gene symbols, Uniprot IDs, Ensembl IDs and gene locations is given as a semicolon-delimited list of entries. In particular, we included all UniProt IDs for a given gene, which is essential for a robust mapping. Finally, using the platform IDs as a key, we collated the association information of the instruments from each study based on the Uniprot IDs (**Supplementary Table 26**).

2 Instrument validation

2.1. Combining and reassigning instruments to proteins

Since we regrouped the protein IDs based on their functions, we further reassigned the instruments to fit the new protein naming system rather than the protein names reported in the pQTL studies. We standardized the format of the instrument files across studies and reassigned them based on their Uniprot ID.

2.2. Instrument specificity

Absence of horizontal pleiotropy is one of the core assumptions for MR, which assumes that the genetic variant should only be related to the outcome of interest through the instrumented exposure. We noted that some SNPs were associated with more than one protein, for example, APOE SNP rs7412 is associated with a set of proteins such as ADAM11, APBB2 and APOB. We considered these instruments associated with more than 5 proteins as potentially pleiotropic SNPs and non-specific for any particular protein level and set up a flag (discrete numeric parameter) based on the number of proteins these SNPs (and their proxies with LD r2>0.5) associated with ("N_protein" column in **Supplementary Table 1**).

2.3. Instrument consistency across studies

2.3.1. Cross-referencing association results between studies

We pooled pQTLs from 5 studies, which have employed different proteomics arrays. The two main assays were the SOMAscan aptamer-based multiplex protein array [22] and the OLINK ProSeek CVD array [23]. The SOMAscan platform is based on the technology called Slow Off-rate Modified Aptamer (SOMAmer), for which reagents consist of a short single-stranded DNA sequence that incorporates a series of modifications that give the SOMAmer "protein-like" appendages. The OLINK ProSeek method is based on the highly sensitive and specific proximity extension array, which involves the binding of distinct polyclonal oligonucleotide-labelled antibodies to the target protein followed by quantification by real-time quantitative PCR. We noted some examples where SNPs were reported to be associated with a protein in one study but did not reach the genome-wide p-value threshold for statistical significance in other studies including the same protein. In these instances, we investigated whether this reflected no statistical evidence of association (in which case, this inconsistency may indicate potentially artefactual associations) or simply fluctuation of association strength with directionally consistent signals in both studies (which would provide supporting evidence for an instrument). Results of the pair-wise comparisons and the number of SNPs included in each comparison can be found in **Supplementary Table 2**.

We noted some examples where SNPs were reported to be associated with a protein in one study but not reached the genome-wide p value threshold in other studies which had measured the same protein. In these instances, we investigated whether this reflected a no statistical evidence of association (in which case, this inconsistency may indicate potentially artefactual associations) or simply fluctuation of association strength, but with directionally consistent and nominally significant ($p<0.05$) associations in both studies (which would provide supporting evidence for an instrument).

Because of the low number of proteins measured in Folkersen *et al* and Yao *et al*, the number of cases where we could perform a validation across 4 or even 3 studies were limited. Instead, for the 1062 pQTLs with SNP lookup results in at least 2 studies, we performed 9 pair-wise comparisons to assess the consistency of the SNP effects of instruments in each pair of studies. Firstly, we tested the overall agreement of effect estimates for the pairwise comparisons. We estimated the pair-wise correlation (r), 95% confidence intervals and p values using the "cor.test" function in R (https://www.r-project.org/).

To provide an overall visualisation of the agreements we generated scatter plots for all pair-wise combinations (**Supplementary Figure 1**). For each scatter plot, we compared the genetic associations (betas, 95% CIs) from one study with the same SNPs looked up in the other study. We also generated Bland-Altman plots to compare the genetic effects between different studies (**Supplementary Figure 2**).

In summary, we show that the agreement of pQTL regression coefficients is high across all five studies (correlation ranged from r = 0.58 to 0.94, **Supplementary Table 2**), and in general, the effects of the SNP associations for all nine study-level pair-wise comparisons follows the identity line well with few outliers (**Supplementary Figure 1 and 2**). We further performed two consistency tests on the instruments which were present across studies.

2.3.2. Pair-wise instrument validation

The first consistency test was a *heterogeneity test* using a pair-wise Z statistic to investigate whether there was statistical evidence of heterogeneity between effect sizes in different studies (for all pQTL studies included in our analysis where: 1) effect sizes were always in SD unit; 2) using similar sets of covariates). If the Z score was greater than 5 (equal to a P value of 0.001), we considered the instrument to have strong evidence of heterogeneity indicating inconsistency of effect sizes between studies.

Furthermore, we defined three flags (binary parameter) based on 1) whether the direction of the effects agreed across studies (column "Agree_beta" in **Supplementary Table 15**); 2) whether P values of the SNP association were smaller than 0.05 across studies (column "Agree_P" in **Supplementary Table 15**); 3) whether there was statistical evidence of heterogeneity between the two effect sizes (pair-wise Z test, with Z score greater than 5 as threshold, which is equal to p value of 0.001) (column "Heterogeneity_across_studies" in **Supplementary Table 1**).

2.3.3. Instrument validation using colocalization analysis across protein studies

The second consistency test was a *colocalization analysis*, which estimates the posterior probability (PP) of the same protein measured in different studies sharing the same causal pQTL within a 1Mb window around the pQTL with the smallest P value. The default priors for colocalization analysis were used here (the prior probability a SNP is associated with the protein is $1x10^{-4}$; the prior probability a SNP is associated with the human phenotype is $1x10^{-4}$; and the prior probability a SNP is associated with both the protein and the phenotype is $1x10^{-5}$). We also applied the pair-wise conditional and colocalization analysis (PWCoCo) for regions with multiple pQTLs to avoid the assumptions of traditional colocalization approaches of just a single association signal per region (details in **Online methods**: *Pair-wise conditional and colocalization analysis*). A lack of evidence (i.e. PP< 80%) in the conventional colocalization and PWCoCo analysis would suggest that the pQTL reported in the two studies did not share the same causal signals within the region, therefore are not consistent between the studies. The colocalization analysis was conducted using the "coloc" R package [24]. For instruments with SNP association information in both Sun *et al* and Folkersen *et al*, we were able to conduct colocalization analysis. However, due to lack of sufficient SNP coverage, it was not possible to conduct colocalization analysis to compare the pQTLs from the Emilsson *et al*, Suhre *et al* and Yao *et al* studies. We therefore conducted a LD check for these pQTLs instead. For proteins measured in multiple studies, we estimated the LD between the sentinel variant for each pQTL from one study and the top 30 associated SNPs of the other study in the same region. For pQTLs that showed only weak LD (r^2 < 0.8) with any of the top 30 associated SNPs in the other study, we considered the pQTLs to not share the same causal SNP in the region and therefore be inconsistent instruments.

Distinguishing vertical and horizontal pleiotropic instruments using biological pathway data

Non-specific instruments may exhibit vertical pleiotropy (pQTL associated with proteins on the same pathway) or horizontal pleiotropy (pQTL associated with proteins on different pathways). Vertical pleiotropy does not violate the "exclusion restriction criterion" of MR but horizontal pleiotropy does [25 26]. For any instrument associated with multiple proteins, if these proteins are mapped to the same biological pathway and/or a protein-protein interaction (PPI) exists between them, then, by definition, the instrument is more likely to act through vertical pleiotropy and it is more likely to be a valid instrument for MR. Consequently, as an approach to distinguish vertical from horizontal pleiotropy, we checked the number of pathways and PPIs each protein is involved in for all the instruments associated with 2 to 5 proteins. We used EpiGraphDB (http://www.epigraphdb.org) to extract the most specific (lowest level) pathway information related to each protein from Reactome [71 72] and high confidence PPIs from StringDB (confidence score >0.7) [73 74]. First, we systematically evaluated the number of pathways each protein is involved in (either directly or as part of a complex), and how many PPIs they have. Note, that although the original databases are curated, we may expect some missing information. We further evaluated how many pathways and PPIs are shared between groups of proteins that are associated with the same SNP or SNPs in strong LD (r^2>0.8). The number of shared components for each group of proteins is presented in **Supplementary Table 1**, and **Supplementary Data 2** depicts a detailed comparison within each group using Venn diagrams.

Directionality tests

With sufficiently large sample sizes, a SNP associated with an outcome through a mediating exposure could reach the conventional threshold for statistical significance in both the outcome and exposure GWAS. Therefore, using such thresholds to define instruments could lead to situations where the instrumental SNP influences the hypothesised exposure via the hypothesised outcome (i.e. the hypothesised outcome actually has a causal effect on the hypothesised exposure and not vice versa).

Reverse Mendelian randomization

For associations between proteins and phenotypes identified in the MR analysis, we applied bi-directional MR to evaluate evidence for causal effects in the reverse direction by modelling complex phenotypes as our exposure and plasma protein as our outcome. Instruments for complex phenotypes were selected based on a threshold of $P < 5 \times 10^{-8}$ from GWAS after LD clumping to identify independent variants. The IVW method was applied to estimate the causal effects of phenotypes on proteins where more than one instrument was available, otherwise the Wald ratio was used. MR-Egger [31] was used as a sensitivity analysis to test for potential pleiotropic effects.

Identifying the direction of effects for instruments using Steiger filtering

Due to lack of sufficient SNP association information (e.g. allele information, effect size, standard error) for some pQTL studies, it was not possible to conduct bi-directional MR using all proteins as outcomes. Therefore, we conducted Steiger filtering as an alternative method to test the directionality of protein-phenotype associations. The Steiger method [32] has been implemented in the TwoSampleMR R package [33] to assess directionality of instrument-phenotype associations [34 35]. For disease phenotypes, we estimated the variance explained on the liability scale. We then set up a flag (categorical variable) to record the direction of the effects of the SNPs using Steiger filtering.

Steiger filtering acts slightly different for MR using cis or trans pQTLs. For cis pQTLs, measurement error may bias the results. For trans pQTLs, a confounder may bias the results. However, the bias from these issues is expected to be minimal.

Linkage disequilibrium check

Results that survived the multiple testing threshold in the MR analysis were evaluated using a stringent Bayesian model (colocalization analysis) to estimate the posterior probability (PP) of each genomic locus containing a single variant affecting both the protein and the phenotype [24]. The default priors were used for the analysis. A PP > 80% in this analysis would suggest that the two association signals are likely to colocalize within the test region. Colocalization analysis is commonly conducted for cis QTLs [36 37] but under studied for trans QTLs. We therefore applied colocalization to both cis and trans pQTLs. For protein and phenotype GWAS lacking sufficient SNP coverage or missing key information (e.g. allele frequency or effect size) in the test region, we conducted a LD check for the sentinel variant for each pQTL against the 30 strongest SNPs in the region associated with the phenotype as an approximate colocalization analysis. r^2 of 0.8 between the sentinel pQTL variant and any of the 30 strongest SNPs associated with the phenotype was used as evidence for approximate colocalization. For all MR top findings, we treated colocalised findings (PP>=80%) as "Colocalised" and LD checked findings (r^2>=0.8) as "LD checked"; other findings that did not pass the colocalization or LD check analysis were annotated as "Not colocalized". For MR findings using multiple instruments (e.g. cis + trans analysis), we tested each pQTL with the phenotype separately. Only if all pQTLs colocalised with the phenotype at r^2>=0.8 did we treat this finding as colocalised.

References

1. Xu, Y. & Krukoff, T. L. Adrenomedullin in the rostral ventrolateral medulla increases arterial pressure and heart rate: roles of glutamate and nitric oxide. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **287**, R729-34 (2004).
2. Shindo, T. *et al.* Vascular abnormalities and elevated blood pressure in mice lacking adrenomedullin gene. *Circulation* **104**, 1964–1971 (2001).
3. Hill, L. D. *et al.* Fetal ERAP2 variation is associated with preeclampsia in African Americans in a case-control study. *BMC Med. Genet.* **12**, 64 (2011).
4. Castelletti, F. *et al.* Mutations in FN1 cause glomerulopathy with fibronectin deposits. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 2538–2543 (2008).
5. Van Buren, P. N. & Toto, R. Hypertension in diabetic nephropathy: epidemiology, mechanisms, and management. *Adv. Chronic Kidney Dis.* **18**, 28–41 (2011).
6. Wuttge, D. M. *et al.* CXCL16/SR-PSOX is an interferon-gamma-regulated chemokine and scavenger receptor expressed in atherosclerotic lesions. *Arterioscler. Thromb. Vasc. Biol.* **24**, 750–755 (2004).
7. Norlander, A. E., Saleh, M. A. & Madhur, M. S. CXCL16: a chemokine-causing chronic kidney disease. *Hypertension* **62**, 1008–1010 (2013).
8. Yeap, B. B. *et al.* IGF1 and its binding proteins 3 and 1 are differentially associated with metabolic syndrome in older men. *Eur. J. Endocrinol.* **162**, 249–257 (2010).
9. Ganesh, S. K. *et al.* Effects of long-term averaging of quantitative blood pressure traits on the detection of genetic associations. *Am. J. Hum. Genet.* **95**, 49–65 (2014).
10. Kurohara, K. *et al.* Essential roles of Meltrin beta (ADAM19) in heart development. *Dev. Biol.* **267**, 14–28 (2004).
11. Lee, K.-Y. *et al.* Inter-alpha-trypsin inhibitor heavy chain 4: a novel biomarker for environmental exposure to particulate air pollution in patients with chronic obstructive pulmonary disease. *Int. J. Chron. Obstruct. Pulmon. Dis.* **10**, 831–841 (2015).

12. Li, Y. & Yang, P. GPC5 gene and its related pathways in lung cancer. *J. Thorac. Oncol.* **6**, 2–5 (2011).
13. Ren, J.-G., Jie, C. & Talbot, C. How PEDF prevents angiogenesis: a hypothesized pathway. *Med. Hypotheses* **64**, 74–78 (2005).
14. Homan, E. P. *et al.* Mutations in SERPINF1 cause osteogenesis imperfecta type VI. *J. Bone Miner. Res.* **26**, 2798–2803 (2011).
15. Soler Artigas, M. *et al.* Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat. Genet.* **43**, 1082–1090 (2011).
16. Bilsborough, J., Targan, S. R. & Snapper, S. B. Therapeutic Targets in Inflammatory Bowel Disease: Current and Future. *The American Journal Of Gastroenterology Supplements* **3**, 27 (2016).
17. Hirota, T. *et al.* Genome-wide association study identifies eight new susceptibility loci for atopic dermatitis in the Japanese population. *Nat. Genet.* **44**, 1222–1226 (2012).
18. Paternoster, L. *et al.* Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. *Nat. Genet.* **47**, 1449–1456 (2015).
19. Lin, X., Zhang, Y. & Chen, Q. FCRL3 gene polymorphisms as risk factors for rheumatoid arthritis. *Hum. Immunol.* **77**, 223–229 (2016).
20. Sun, B. B. *et al.* Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018).
21. Folkersen, L. *et al.* Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular disease. *PLoS Genet.* **13**, e1006706 (2017).
22. Rohloff, J. C. *et al.* Nucleic Acid Ligands With Protein-like Side Chains: Modified Aptamers and Their Use as Diagnostic and Therapeutic Agents. *Mol. Ther. Nucleic Acids* **3**, e201 (2014).
23. Assarsson, E. *et al.* Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS One* **9**, e95192 (2014).
24. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).

25. Hemani, G., Bowden, J. & Davey Smith, G. Evaluating the potential role of pleiotropy in Mendelian randomization studies. *Hum. Mol. Genet.* **27**, R195–R208 (2018).
26. Swerdlow, D. I. *et al.* Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. *Int. J. Epidemiol.* **45**, 1600–1616 (2016).
27. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).
28. Fabregat, A. *et al.* Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinformatics* **18**, 142 (2017).
29. Szklarczyk, D. *et al.* The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).
30. Szklarczyk, D. *et al.* STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447-52 (2015).
31. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
32. Steiger, J. H. Tests for comparing elements of a correlation matrix. *Psychol. Bull.* **87**, 245 (1980).
33. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**, (2018).
34. Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* **13**, e1007081 (2017).
35. Hemani, G. *et al.* Automating Mendelian randomization through machine learning to construct a putative causal map of the human phenome. bioRxiv (2017) doi:10.1101/173682.
36. Zhu, Z. *et al.* Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).

37. Richardson, T. G. *et al.* Systematic Mendelian randomization framework elucidates hundreds of CpG sites which may mediate the influence of genetic variants on disease. *Hum. Mol. Genet.* (2018) doi:10.1093/hmg/ddy210.